A GOOD UNIVERSAL WEIGHT FOR MULTIPLE RECURRENCE AVERAGES WITH COMMUTING TRANSFORMATIONS IN NORM

IDRIS ASSANI AND RYO MOORE

Abstract. We will show that the sequences appearing in Bourgain's double recurrence result are good universal weights to the multiple recurrence averages with commuting measure-preserving transformations in norm. This will extend the pointwise converge result of Bourgain, the norm convergence result of Tao, and the authors' previous work on the single measure-preserving transformation. The proof will use the double-recurrence Wiener-Wintner theorem, factor decompositions (Host-Kra-Ziegler factors), nilsequences, and various seminorms including the ones by Gowers-Host-Kra as well as the box seminorms introduced by Host.

1. Introduction

In this paper, which is a sequel to [7], we will continue the study of the return times theorem and its connection to multiple ergodic recurrence. Such study was initiated by the first author in 2000 [2]. More details on the historical background on the return times theorem can be found in the survey paper prepared by the first author and Presser [8]. The same notations are used in the previous paper [7].

1.1. **Process and good universal weights.** We recall that $(X_n)_n$ is a *process* if for all nonnegative integers n, X_n is a bounded and measurable function on some probability space $(\Omega, \mathcal{S}, \mathbb{P})$ (cf. [7, Definition 1.1]). Here we give a slightly more precise definition of the good universal weights.

Definition 1.1. *We denote by*

$$M_1 = \left\{ (a_n) : \sup_N \frac{1}{N} \sum_{n=1}^{N} |a_n| < \infty \right\}.$$

We denote Π to be a collection of probability measure spaces, and $\mathfrak{X}(\Pi)$ be a collection of processes for some probability measure space $(\Omega, \mathcal{S}, \mathbb{P}) \in \Pi$, i.e.

$$(X_n) \in \mathfrak{X}(\Pi) \implies \textit{for all } n \geq 1, X_n : \Omega \to \mathbb{C} \textit{ bounded and measurable on some } (\Omega, \mathcal{S}, \mathbb{P}) \in \Pi.$$

- *We say a sequence $(a_n) \in M_1$ is a* ***good universal weight for*** $\mathfrak{X}(\Pi)$ ***(a.e.) pointwise****, if for any probability space $(\Omega, \mathcal{S}, \mathbb{P}) \in \Pi$, and any process $(X_n) \in \mathfrak{X}(\Pi)$ on Ω, the averages*

$$\frac{1}{N} \sum_{n=1}^{N} a_n X_n(\omega)$$

converge for $\mathbb{P}$-a.e. $\omega \in \Omega$.

- *We say a sequence* $(a_n) \in M_1$ *is a* ***good universal weight for*** $\mathfrak{X}(\Pi)$ ***in norm****, if for any probability space* $(\Omega, \mathcal{S}, \mathbb{P}) \in \Pi$ *and any process* $(X_n) \in \mathfrak{X}(\Pi)$ *on* Ω*, the averages*
$$\frac{1}{N}\sum_{n=1}^{N} a_n X_n(\omega)$$
converge in $L^2(\mathbb{P})$

For instance, Bourgain's return times theorem [10,12] can be stated as follows: Given an ergodic system $(X, \mathcal{F}, \mu, T)$ and $f \in L^\infty(\mu)$, for μ-a.e. $x \in X$, the sequence $(f(T^n x))_n$ is a good universal weight for $\mathfrak{X}(\Pi)$ pointwise, where Π is the collection of all the measure-preserving system, and
$$\mathfrak{X}(\Pi) = \{g \circ S^n : (Y, \mathcal{G}, \nu, S) \in \Pi, g \in L^\infty(\nu)\}.$$
Also, the result from our previous paper [7] can be said as follows: Given a measure-preserving $(X, \mathcal{F}, \mu, T)$, functions $f_1, f_2 \in L^\infty(\mu)$, and any distinct integers a and b, the sequence $(f_1(T^{an}x) f_2(T^{bn}x))$ is μ-a.e. a good universal weight for $\mathfrak{X}(\Pi)$ in norm, where Π is a collection of probability measure-preserving systems, and
$$\mathfrak{X}(\Pi) = \left\{\prod_{i=1}^{k} g_i \circ S^{ik} : (Y, \mathcal{G}, \nu, S) \in \Pi, g_1, \dots, g_k \in L^\infty(\nu)\right\}.$$
We shall call this class of processes $\mathfrak{X}(\Pi)$ the *linear multiple recurrence averages with single transformation*.

We note that this result extends the work of Host and Kra from 2009 in [21, Theorem 2.25], where they had $a = 1$ and $f_2 = \mathbb{1}_X$. We also remark here that Host and Kra obtained their result as a consequence of the nilsequence Wiener-Wintner averages [21, Theorem 2.22], whereas ours only relied on the double recurrence (classical) Wiener-Wintner averages. We recall that the study of the double recurrence Wiener-Wintner theorem was initiated by Duncan (the first author's former Ph.D. student) in his doctoral dissertation completed in 2001 [14]. The result was later generalized by himself and the authors in 2014 [5], and further extended to a polynomial Wiener-Wintner by us in [6]. Later, the first author showed that the double recurrence nilsequence Wiener-Wintner averages converge off a single null set [4], using the techniques that can be seen in his work of averages along cubes [3], the paper by Host and Kra [21], and on the work of the classical double recurrence Wiener-Wintner result [5]. This answered B. Weiss's question that was asked during the 2014 Ergodic Theory Workshop at the UNC-Chapel Hill positively.[1]

We remark that P. Zorin-Kranich mentioned in [27, v.2, Proposition 1.3] that if a bounded sequence (c_n) is a good universal weight for multiple ergodic averages with a single transformation, then a proposition suggested by Frantzikinakis [16, Proposition 2.4] says that (c_n) satisfies the nilsequence Wiener-Wintner averages. We recall that, however, Frantzikinakis suggests in the remark of [16, Proposition 2.4] that one is

[1]The double recurrence nilsequence Wiener-Wintner result was announced during the second author's Ph.D. oral exam on April 10th, 2015, and the preprint of this result was submitted to arXiv.org on Wednesday, April 22nd, 2015 (11:12:30 GMT)—the day after the first version of a preprint announcing that the similar result was submitted to arXiv.org by P. Zorin-Kranich [27], which was submitted on Friday, April 17, 2015 (22:06:28 GMT) that ultimately appeared on arXiv.org on Tuesday, April 21, 2015 (0 GMT). Zorin-Kranich posted the second version on arXiv.org on August 5th, 2015.

required to show a "nontrivial variant" of a work of Green and Tao [18, Lemma 14.2]. If, however, (c_n) is a sequence appearing in the double recurrence averages (i.e. $c_n = f_1(T^{an}x)f_2(T^{bn}x)$), the following holds:

Proposition 1.2. *Let $(X, \mathcal{F}, \mu, T)$ be a measure-preserving system, $f_1, f_2 \in L^\infty(\mu)$, and a, b be distinct integers. The following statements are equivalent.*

(i) *The sequence $(f_1(T^{an}x)f_2(T^{bn}x))_n$ is a good universal weight for the linear multiple recurrence averages with single transformation in norm.*

(ii) *The classical double recurrence Wiener-Wintner averages converge off a single null set, i.e. there exists a set of full measure X_{f_1,f_2} such that for any $x \in X_{f_1,f_2}$ the averages*

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)e^{2\pi int}$$

converge for any $t \in \mathbb{R}$.

(iii) *The nilsequence Wiener-Wintner averages converge off a single null set, i.e. there exists a set of full measure X'_{f_1,f_2} such that for any $x \in X'_{f_1,f_2}$ and for any nilsequence (b_n), the averages*

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T^{an}x)f_2(T^{bn}x)b_n$$

converge.

The proof of this proposition is given in the revised version of [4] following referee comments.

1.2. **The main theorem.** In the series of work on convergence of multiple recurrent averages, Tao [23] showed that given any measure-preserving system with multiple commuting transformations $(Y, \mathcal{G}, \nu, S_1, S_2, \ldots, S_k)$ and any functions $g_1, g_2, \ldots, g_k$, the averages

$$\frac{1}{N}\sum_{n=0}^{N-1}\prod_{i=1}^{k} g_i \circ S_i^n$$

converge in $L^2(\nu)$. Followed by his result, different proofs were obtained by Austin [9], Host [19], and Towsner [25]. For the pointwise convergence, on the other hand, Bourgain [11] showed in 1990 that the case holds for $k = 2$ and S_i is a different nonzero power of a measure-preserving transformation S. With some more assumption on the space, the first author [1, Theorem 2] showed in 1998 that the pointwise convergence holds for any k, where again each S_i is a different nonzero power of a measure-preserving transformation S, providing examples of the good universal weights discussed in the return times result from 2000 [2, Theorem 3]. Recently, Huang, Shao, and Ye announced the pointwise convergence of the linear multiple ergodic averages with single transformation for the case of distal system [22]. In this paper, we will show that the double recurrence sequence that appeared in the work of Bourgain is a good universal weight for the multiple recurrence averages with commuting transformations in L^2-norm. This will extend the double recurrence result of Bourgain and the norm convergence result of Tao simultaneously, although both of them are assumed in the arguments of this paper.

Theorem 1.3 (The main result)**.** *Let $(X, \mathcal{F}, \mu, T)$ be a measure-preserving system, and suppose $f_1, f_2 \in L^\infty(\mu)$. Then there exists a set of full measure X_{f_1,f_2} such that for any $x \in X_{f_1,f_2}$, for any $a, b \in \mathbb{Z}$ and any positive integer $k \geq 1$, for any other measure-preserving system with k commuting transformations $(Y, \mathcal{G}, \nu, S_1, S_2, \ldots S_k)$ for any $k \in \mathbb{N}$, and for any $g_1, g_2, \ldots g_k \in L^\infty(\nu)$, the averages*

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^{an}x) f_2(T^{bn}x) \prod_{i=1}^{k} g_i \circ S_i^n \text{ converge in } L^2(\nu). \tag{1}$$

We note that this theorem generalized our previous work, where each transformation on Y is a different power of the first one (i.e. $S_i = S^i$ for $1 \leq i \leq k$) [7]. The extension to the commuting case considered in this paper is not a immediate consequence of the previous result. However, the general framework created for tackling the single transformation case is shown to be useful to prove Theorem 1.3, i.e.

Step I. Use the uniform Wiener-Wintner theorem for the double recurrence [5] and the spectral theorem to show the the averages converge to zero in norm by induction, provided that f_1 and f_2 belong to the orthogonal complement of an appropriate Host-Kra-Ziegler factor.

Step II. When f_1 and f_2 are measurable with respect to the appropriate Host-Kra-Ziegler factor, we obtain the norm convergence for the case using the structure of nilsystems [20].

In the previous result [7], we were able to use the characteristic factor of the other system to obtain the result. This is no longer ideal for the proof of Theorem 1.3, as the method of using characteristic factor involving multiple transformations seems unrealistic, as the difficulty is suggested by Host [19]. Instead of relying on the structure of the other system, we approximated the sequence $(f_1(T^{an}x)f_2(T^{bn}x))$ by a nilsequence with vertical frequency, and utilized the box seminorms and magic systems that were introduced by Host [19] to obtain the convergence result.

In terms of Definition 1.1, Theorem 1.3 states that for μ-a.e. $x \in X$, the sequence $(f_1(T^{an}x)f_2(T^{bn}x))_n$ is a good universal weight for $\mathfrak{X}(\Pi)$ in norm, where Π is a collection of measure-preserving systems with multiple commuting transformations, and

$$\mathfrak{X}(\Pi) = \left\{ \prod_{i=1}^{k} g_i \circ S_i^n : (Y, \mathcal{G}, \nu, S_1, \ldots, S_k) \in \Pi, g_1, g_2, \ldots, g_k \in L^\infty(\nu) \right\}.$$

Throughout this paper, we will assume that the system $(X, \mathcal{F}, \mu, T)$ is ergodic, and the result holds for general measure-preserving system after we apply an ergodic decomposition. In the proof of the theorem, we will first consider the case where either f_1 or f_2 belongs to the orthogonal complement of the $k+1$-th Host-Kra-Ziegler factor [20, 26]. For that case, we will show that the averages converge to zero.

Theorem 1.4. *Let notations be as in Theorem 1.3. Suppose that T is ergodic. If either f_1 or f_2 belongs to the orthogonal complement of the $k+1$-th Host-Kra-Ziegler factor of T, then there exists a set of full measure $X^1_{f_1,f_2}$*

such that for any $x \in X^1_{f_1,f_2}$, *we have*

$$\limsup_{N\to\infty} \left\| \frac{1}{N}\sum_{n=0}^{N-1} f_1(T^{an}x) f_2(T^{bn}x) \prod_{i=1}^{k} g_i \circ S_i^k \right\|_{L^2(\nu)} = 0 \tag{2}$$

Next, we will assume that both f_1 and f_2 belong to the $k+1$-th Host-Kra-Ziegler factor. In this case, the sequence $a_n = f_1(T^{an}x)f_2(T^{bn}x)$ can be approximated by a $k+1$-step nilsequence. Thus, the following estimate will be useful.

Theorem 1.5. *Suppose a_n is a $k+1$-step nilsequence for $k \geq 2$. Then*

$$\limsup_{N\to\infty} \left\| \frac{1}{N}\sum_{n=0}^{N-1} a_n \prod_{i=1}^{k} g_i \circ S_i^k \right\|_{L^2(\nu)} \leq [[[g_1]]]_{1,k} \tag{3}$$

where $[[[\cdot]]]_{i,k}$ denotes the box seminorm $|||\cdot|||$ on $L^\infty(\nu)$ (cf. §2.2) that corresponds to the transformations

$$\underbrace{S_i, S_i, \ldots, S_i}_{k+1 \text{ times}}, S_1S_i^{-1}, S_2S_i^{-1}, \ldots, S_{i-1}S_i^{-1}, S_{i+1}S_i^{-1}, \ldots, S_kS_1^{-1}.$$

Throughout this paper, we will assume that the functions appearing (such as f_i's, g_j's) are real-valued, and will assume that $|f_i| \leq 1$ and $|g_j| \leq 1$ for $i = 1,2$ and $j = 1,2,\ldots,k$.

2. Preliminaries

In this section, we will provide a brief summary of results and notations that will be used in our arguments.

2.1. **Host-Kra-Ziegler factors, nilsystems, and nilsequences.** Let $(X, \mathcal{F}, \mu, T)$ be an ergodic system. We will denote $(Z_l, \mathcal{Z}_l, \mu_l, T_l)$ to be the l-th *Host-Kra-Ziegler factor* (cf. [20,26]) of $(X, \mathcal{F}, \mu, T)$. Unless there is a confusion, we will denote μ and T in place of μ_l and T_l.

The Gowers-Host-Kra seminorms (cf. [17,20]) will be denoted as $|||\cdot|||_{l+1}$. It was shown in [20, Lemma 4.3] that if $f \in L^\infty(\mu)$, $|||f|||_{l+1} = 0$ if and only if $\mathbb{E}(f|\mathcal{Z}_l(T)) = 0$.

Let G be a nilpotent Lie group of order l, and Γ be a discrete cocompact subgroup of G. The homogeneous space G/Γ is a *nilmanifold* of order l. Let $N = G/\Gamma$, ρ be the Haar measure on X, $u \in N$, and $U : X \to X$ be the transformation $Ux = u \cdot x$. Then the system (N, ρ, U) is called *nilsystem* of order l. It was shown in [20, Theorem 10.1] that every l-th order Host-Kra-Ziegler factor is an inverse limit of l-th order nilsystems.

Suppose $N = G/\Gamma$ is an l-th order nilsystem, and $\tau \in G$. If $\phi \in \mathcal{C}(N)$, we say $(\phi(\tau^n x))_n$ is a *basic l-step nilsequence* for any $x \in N$. An *l-step nilsequence* is a uniform limit of basic l-step nilsequences.

2.2. **Box measures and seminorms, magic systems.** We also recall the box measures, box seminorms, and the magic systems that were introduced by Host in [19], which he used to provide a different proof to Tao's norm convergence result for commuting transformations [23]. Suppose $(Y, \nu, S_1, S_2, \ldots, S_k)$ is a

system for which $S_1, S_2, \ldots, S_k$ are measure-preserving transformations that commute with each other. We denote $\mathcal{I}(S_i)$ to be the σ-algebra of S_i-invariant sets in Y. We define a *conditionally independent square* $\nu_{S_i} = \nu \times_{\mathcal{I}(S_i)} \nu$ over $\mathcal{I}(S_i)$ to be a measure on Y^2 such that if $g, g' \in L^\infty(\nu)$, we have

$$\int g(y)g'(y')\, d\nu \times_{\mathcal{I}(S_i)} \nu(y, y') = \int \mathbb{E}_\nu(g|\mathcal{I}(S_i))(y)\mathbb{E}_\nu(g'|\mathcal{I}(S_i))(y)\, d\nu(y).$$

Similarly, we can define a measure on Y^4 by letting $\nu_{S_i,S_j} = \nu_{S_i} \times_{\mathcal{I}(S_j \times S_j)} \nu_{S_i}$, where for any $g_\epsilon \in L^\infty(\nu)$, where $\epsilon \in \{0,1\}^2$, we have

$$\int \prod_{\epsilon \in \{0,1\}^2} g_\epsilon(y_\epsilon) d\nu_{S_i,S_j} = \int \mathbb{E}_{\nu_{S_i}}(g_{00} \otimes g_{10}|\mathcal{I}(S_j \times S_j))(y_0, y_1)\mathbb{E}_{\nu_{S_i}}(g_{01} \otimes g_{11}|\mathcal{I}(S_j \times S_j))(y_0, y_1) d\nu_{S_i}(y_0, y_1).$$

By iterating this process, we can define a measure $\nu_{S_1,S_2,\ldots,S_d}$ on Y^{2^d} for $1 \le d \le k$ so that for any $g_\epsilon \in L^\infty(\nu)$ such that $\epsilon \in \{0,1\}^d$, we have

$$\int \prod_{\epsilon \in \{0,1\}^d} g_\epsilon(y_\epsilon) d\nu_{S_1,S_2,\ldots,S_d}$$
$$= \int \mathbb{E}_{\nu_{S_1,\ldots,S_{d-1}}}\left(\bigotimes_{\eta \in \{0,1\}^{d-1}} g_{\eta 0} \,\middle|\, \mathcal{I}(\underbrace{S_d \times \cdots \times S_d}_{2^{d-1} \text{ times}}) \right) \mathbb{E}_{\nu_{S_1,\ldots,S_{d-1}}}\left(\bigotimes_{\eta \in \{0,1\}^{d-1}} g_{\eta 1} \,\middle|\, \mathcal{I}(\underbrace{S_d \times \cdots \times S_d}_{2^{d-1} \text{ times}}) \right) d\nu_{S_1,\ldots,S_{d-1}}.$$

When $d = k$, we will denote the space $Y^{2^k} = Y^*$ and $\nu_{S_1,S_2,\ldots,S_k} = \nu^*$. We say that ν^* is the *box measure associated to the transformations* $S_1, S_2, \ldots, S_k$. On the measure space (Y^*, ν^*), we define *side transformations* S_i^* for $1 \le i \le k$ in the following way:

$$\text{For every } \epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_k) \in \{0,1\}^k,\ (S_i^* y)_\epsilon = \begin{cases} S_i y_\epsilon & \text{if } \epsilon_i = 0, \\ y_\epsilon & \text{if } \epsilon_i = 1. \end{cases}$$

For example, for the case $k = 2$, we have

$$S_1^* = S_1 \times \mathrm{Id} \times S_1 \times \mathrm{Id}, \text{ and } S_2^* = S_2 \times S_2 \times \mathrm{Id} \times \mathrm{Id},$$

and for $k = 3$, we would have

$$S_1^* = S_1 \times \mathrm{Id} \times S_1 \times \mathrm{Id} \times S_1 \times \mathrm{Id} \times S_1 \times \mathrm{Id},$$
$$S_2^* = S_2 \times S_2 \times \mathrm{Id} \times \mathrm{Id} \times S_2 \times S_2 \times \mathrm{Id} \times \mathrm{Id}, \text{ and}$$
$$S_3^* = S_3 \times S_3 \times S_3 \times S_3 \times \mathrm{Id} \times \mathrm{Id} \times \mathrm{Id} \times \mathrm{Id}.$$

Note that the measure ν^* is invariant under each side transformation S_i^* for $1 \le i \le k$, and each S_i^* commute with each other. Hence, $(Y^*, \nu^*, S_1^*, \ldots, S_k^*)$ is a measure-preserving system with k commuting transformations.

Suppose $y^* = (y_\epsilon)_{\epsilon\in\{0,1\}^k} \in Y^*$, and $y_\varnothing$ is the $\varnothing = (0,0,\dots,0) \in \{0,1\}^k$ coordinate entry of y^*. We note that the projection map $\pi : Y^* \to Y$ for which $\pi(y^*) = y_\varnothing$ is a factor map from $(Y^*, \nu^*, S_1^*, \dots, S_k^*)$ to $(Y, \nu, S_1, \dots, S_k)$ (since $\pi \circ S_i^* = S_i \circ \pi$ for each $i = 1, 2, \dots, k$).

We can now define seminorms on $L^\infty(\nu)$ associated to these transformations: For $g \in L^\infty(\nu)$, we define

$$|||g||| = |||g|||_{S_1,S_2,\dots,S_k} := \left(\int \prod_{\epsilon\in\{0,1\}^k} g(y_\epsilon) d\nu^*(y^*) \right)^{1/2^k}.$$

By [19, Proposition 2], $|||\cdot|||$ is indeed a seminorm. Furthermore, we know from [19, Equation (11)] that for every $g \in L^\infty(\nu)$, we have

$$|||g|||^{2^d}_{S_1,\dots,S_d} = \lim_{N_d\to\infty} \frac{1}{N_d} \sum_{n_d=0}^{N_d-1} |||g \cdot g \circ S_d^{n_d}|||^{2^{d-1}}_{S_1,\dots,S_{d-1}}. \tag{4}$$

By the construction of the box seminorms and measures, we know that

$$|||g|||_{S_1,\dots,S_i,\dots,S_k} = |||g|||_{S_1,\dots,S_i^{-1},\dots,S_k} \text{ for any } 1 \le i \le d. \tag{5}$$

It was also shown in [19, Corollary 3] that the box seminorm remains unchanged if the transformations $S_1, S_2, \dots, S_d$ are permuted. Furthermore, by [19, Proposition 1], we have the following estimate: If we denote $T_1 = S_1$, and $T_i = S_i S_1^{-1}$ for $2 \le i \le k$, then

$$\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} \prod_{i=1}^{k} g_i \circ S_i^k \right\|_{L^2(\nu)} \le |||g_1|||_{T_1,T_2,\dots,T_k}. \tag{6}$$

We distinguish these seminorms and the Gowers-Host-Kra seminorms by dropping the numerical subscript to the former.

Let $\mathcal{W}$ be the join of the σ-algebras $\mathcal{I}(S_i)$ for each $i = 1, 2, \dots, k$, i.e.

$$\mathcal{W} = \bigvee_{i=1}^{k} \mathcal{I}(S_i).$$

We say that the system $(Y, \nu, S_1, \dots, S_k)$ is *magic* if the following holds: Given $g \in L^\infty(\nu)$,

$$\mathbb{E}_\nu(g|\mathcal{W}) = 0 \text{ implies that } |||g|||_{S_1,S_2,\dots,S_k} = 0.$$

It was shown in [19, Theorem 2] that $(Y^*, \nu^*, S_1^*, \dots, S_k^*)$ is a magic system, i.e. given $G \in L^\infty(\nu^*)$,

$$\mathbb{E}_{\nu^*}(G|\mathcal{W}^*) = 0 \text{ implies that } |||G|||_{S_1^*,S_2^*,\dots,S_k^*} = 0 \text{ where } \mathcal{W}^* = \bigvee_{i=1}^{k} \mathcal{I}(S_i^*).$$

3. Proof of Theorem 1.4

The proof presented here is analogous to that of the proof of [7, Theorem 1.5(a)][2] for the case we had a single measure-preserving transformation S (i.e. $S_i = S^i$). We recall the following inequality that was

[2]In fact, more details to the proof, including specific cases $k = 2$ and $k = 3$, are presented in the cited reference.

obtained in the proof of the double recurrence Wiener-Wintner result [5]:

$$\int \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=0}^{N-1} f_1(T^{an}x) f_2(T^{bn}x) e^{2\pi i n t} \right|^2 d\mu(x) \lesssim_{a,b} \min_{i=1,2} |||f_i|||_3^2. \tag{7}$$

In this section, we will denote $a_1 = a$ and $a_2 = b$. Furthermore, we will use the following notations in our arguments.

$$\begin{aligned}
&F_{1,\vec{h}(1)} = f_1 \cdot f_1 \circ T^{a_1 h_1}, && F_{2,\vec{h}(1)} = f_2 \cdot f_2 \circ T^{a_2 h_1},\\
&F_{1,\vec{h}(2)} = F_{1,\vec{h}(1)} \cdot F_{1,\vec{h}(1)} \circ T^{a_1 h_2}, && F_{2,\vec{h}(2)} = F_{2,\vec{h}(1)} \cdot F_{2,\vec{h}(1)} \circ T^{a_2 h_2},\\
&\cdots, && \cdots,\\
&F_{1,\vec{h}(k-1)} = F_{1,\vec{h}(k-2)} \cdot F_{1,\vec{h}(k-2)} \circ T^{a_1 h_{k-1}}, && F_{2,\vec{h}(k-1)} = F_{2,\vec{h}(k-2)} \cdot F_{2,\vec{h}(k-2)} \circ T^{a_2 h_{k-1}}.
\end{aligned}$$

Lemma 3.1. *Let all the notations be as above. Then for each positive integer $k \geq 2$, we have*

$$\begin{aligned}
&\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=1}^{N} f_1(T^{a_1 n}x) f_2(T^{a_2 n}x) \prod_{i=1}^{k} g_i \circ S_i^n \right\|_{L^2(\nu)}^2 \\
&\lesssim_{a_1,a_2} \liminf_{H_1\to\infty} \Bigg(\frac{1}{H_1} \sum_{h_1=1}^{H_1} \liminf_{H_2\to\infty} \frac{1}{H_2} \sum_{h_2=1}^{H_2} \cdots \\
&\liminf_{H_{k-1}\to\infty} \frac{1}{H_{k-1}} \sum_{h_{k-1}=1}^{H_{k-1}} \limsup_{N\to\infty} \sup_{t\in\mathbb{R}} \left| \frac{1}{N} \sum_{n=1}^{N} F_{1,\vec{h}(1)}(T^{a_1 n}x) F_{2,\vec{h}(1)}(T^{a_2 n}x) e^{2\pi i n t} \right|^2 \Bigg)^{2^{-(k-1)}}.
\end{aligned} \tag{8}$$

Proof. We will show this by induction on k. The prove the base case $k = 2$, we first apply van der Corput's lemma to see that

$$\begin{aligned}
&\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} f_1(T^{a_1 n}x) f_2(T^{a_2 n}x) g_1(S_1^n y) g_2(S_2^n y) \right\|_{L^2(\nu)}^2 \\
&\lesssim \liminf_{H_1\to\infty} \frac{1}{H_1} \sum_{h_1=0}^{H_1-1} \int \left| (g_1 \cdot g_1 \circ S_1^h)(y) \limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} F_{1,h_1}(T^{an}x) F_{2,h_1}(T^{bn}x) (g_2 \cdot g_2 \circ S_2^h)((S_2 S_1^{-1})^n y) \right| d\nu.
\end{aligned}$$

By Hölder's inequality (and recalling that $\|g_1\|_{L^\infty(\nu)} \leq 1$), we dominate the last line above by

$$\liminf_{H_1\to\infty} \frac{1}{H_1} \sum_{h_1=0}^{H_1-1} \left(\int \limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=0}^{N-1} F_{1,h_1}(T^{a_1 n}x) F_{2,h_1}(T^{a_2 n}x) (g_2 \cdot g_2 \circ S_2^h)((S_2 S_1^{-1})^n y) \right|^2 d\nu \right)^{1/2}.$$

Let $\sigma_{g\cdot g\circ S_2^h}$ be the spectral measure of $\mathbb{T}$ for the function $g \cdot g \circ S_2^h$ for each h, with respect to the transformation $S_2 S_1^{-1}$. By the spectral theorem, the last expression becomes

$$\liminf_{H_1\to\infty} \frac{1}{H_1} \sum_{h_1=0}^{H_1-1} \left(\int \limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=0}^{N-1} F_{1,h_1}(T^{a_1 n}x) F_{2,h_1}(T^{a_2 n}x) e(nt) \right|^2 d\sigma_{g_2\cdot g_2\circ S_2^h}(t) \right)^{1/2},$$

which is bounded above by

$$\liminf_{H_1\to\infty}\frac{1}{H_1}\sum_{h_1=0}^{H_1-1}\left(\limsup_{N\to\infty}\sup_{t\in\mathbb{R}}\left|\frac{1}{N}\sum_{n=0}^{N-1}F_{1,h_1}(T^{a_1 n}x)F_{2,h_1}(T^{a_2 n}x)e(nt)\right|^2\right)^{1/2}.$$

After we apply the Cauchy-Schwarz inequality (on the averages over H_1), we obtained the desired inequality for the case $k=2$.

Now suppose the estimate holds when we have $k-1$ terms. By applying van der Corput's lemma and the Cauchy-Schwarz inequality, the left hand side of the estimate (8) is bounded above by the product of a constant that only depends on the values of a_1 and a_2 and

$$\liminf_{H_1\to\infty}\left(\frac{1}{H_1}\sum_{h_1=1}^{H_1}\limsup_{N\to\infty}\left\|\frac{1}{N}\sum_{n=1}^{N}F_{1,\vec{h}(k-1)}(T^{a_1 n}x)F_{2,\vec{h}(k-1)}(T^{a_2 n}x)\prod_{i=2}^{k}(g_i\cdot g_i\circ S_i^{h_1})\circ(S_iS_1^{-1})^n\right\|_{L^2(\nu)}^2\right)^{1/2},$$

and we can apply the inductive hypothesis on this lim sup of the square of the L^2-norm above and the Cauchy-Schwarz inequality to obtain the desired estimate. □

The preceding lemma allows us to identify the desired set of full measure for each positive integer k.

Proof of Theorem 1.4. We will first show that for each positive integer $k\geq 1$, there exists a set of full measure $\tilde{X}_k$ such that the statement of Theorem 1.4 holds for this particular k.

The set $\tilde{X}_1$ can be obtained from the double recurrence Wiener-Wintner result [5] by applying the spectral theorem. For $k\geq 2$, we consider a set

$$\tilde{X}_k=\Bigg\{x\in X:\liminf_{H_1\to\infty}\Bigg(\frac{1}{H_1}\sum_{h_1=1}^{H_1}\liminf_{H_2\to\infty}\frac{1}{H_2}\sum_{h_2=1}^{H_2}\cdots$$
$$\liminf_{H_k\to\infty}\frac{1}{H_{k-1}}\sum_{h_k=1}^{H_k}\limsup_{N\to\infty}\sup_{t\in\mathbb{R}}\left|\frac{1}{N}\sum_{n=1}^{N}F_{1,\vec{h}(k-1)}(T^{a_1 n}x)F_{2,\vec{h}(k-1)}(T^{a_2 n}x)e^{2\pi i n t}\right|^2\Bigg)^{2^{-(k-1)}}=0\Bigg\}.$$

We will show that the set on the right hand side is indeed the desired set of full measure. To first show that $\mu(\tilde{X}_k)=1$, we compute that

$$\int\liminf_{H_1\to\infty}\Bigg(\frac{1}{H_1}\sum_{h_1=1}^{H_1}\liminf_{H_2\to\infty}\frac{1}{H_2}\sum_{h_2=1}^{H_2}\cdots \tag{9}$$
$$\liminf_{H_{k-1}\to\infty}\frac{1}{H_{k-1}}\sum_{h_{k-1}=1}^{H_{k-1}}\limsup_{N\to\infty}\sup_{t\in\mathbb{R}}\left|\frac{1}{N}\sum_{n=1}^{N}F_{1,\vec{h}(k-1)}(T^{a_1 n}x)F_{2,\vec{h}(k-1)}(T^{a_2 n}x)e^{2\pi i n t}\right|^2\Bigg)^{2^{-(k-1)}}d\mu=0,$$

which would show that the non-negative term inside the integral equals zero for μ-a.e. $x\in X$. To do so, we apply Fatou's lemma and Hölder's inequality to show that the integral above is bounded above by

$$\liminf_{H_1\to\infty}\Bigg(\frac{1}{H_1}\sum_{h_1=1}^{H_1}\liminf_{H_2\to\infty}\frac{1}{H_2}\sum_{h_2=1}^{H_2}\cdots$$

$$\liminf_{H_{k-1}\to\infty}\frac{1}{H_{k-1}}\sum_{h_{k-1}=1}^{H_{k-1}}\int \limsup_{N\to\infty}\sup_{t\in\mathbb{R}}\left|\frac{1}{N}\sum_{n=1}^{N}F_{1,\vec{h}(k-1)}(T^{a_1 n}x)F_{2,\vec{h}(k-1)}(T^{a_2 n}x)e^{2\pi i n t}\right|^2 d\mu\Bigg)^{2^{-(k-1)}}.$$

Note that the last integral is bounded above by $C\cdot\min_{i=1,2}\left|\left|\left|F_{i,\vec{h}(k-1)}\right|\right|\right|_3^2$ by the estimate (7), where C is a constant that only depends on a_1 and a_2. By letting H_j go to infinity for each $j=1,2,\dots,k-1$, we conclude that the integral on the left hand side of (9) is bounded above by C times the minimum of the power of $|||f_1|||_{k+2}$ or $|||f_2|||_{k+2}$. Since either f_1 or f_2 belongs to $\mathcal{Z}_{k+1}(T)^{\perp}$, we know that either $|||f_1|||_{k+2}=0$ or $|||f_2|||_{k+2}=0$. Thus, (9) holds, which implies that $\tilde{X}_k$ is indeed a set of full measure.

Now we need to show that if $x\in\tilde{X}_k$, then (2) holds. But this follows immediately from Lemma 3.1, since if $x\in\tilde{X}_k$, the right hand side of (8), which is an upper bound for the lim sup of the averages in (2), is 0.

Hence, we conclude the proof by setting $X^1_{f_1,f_2}=\bigcap_{k=1}^{\infty}\tilde{X}_k$. We note that $X^1_{f_1,f_2}$ is a countable intersection of sets of full measures, so $X^1_{f_1,f_2}$ must be a set of full measure as well. □

4. Proof of Theorem 1.5

In this section, we will consider the case where both f_1 and f_2 are measurable with respect to $\mathcal{Z}_{k+1}(T)$. If $(Z_{k+1},\mathcal{Z}_{k+1}(T),\mu,T)$ is the $(k+1)$-th Host-Kra-Ziegler factor, then [20, Theorem 10.1] tells us that it is an inverse limit of nilsystems of order $k+1$. Hence, we can approximate the sequence $(f_1(T^{an}x)f_2(T^{bn}x))_n$ by a $k+1$-step nilsequence, which we shall denote (a_n). We further assume that this nilsequence (a_n) has vertical frequency so that when we apply a multiplicative derivative (as when we use van der Corput's lemma) of an l-step nilsequence $\overline{a}_n a_{n+h}$ is an $l-1$-step nilsequence for any $h\in\mathbb{Z}$ (cf. [15, p. 3505] or [24, Lemma 1.6.13]). Because a set of the linear combination of $\le l$-step nilsequences with vertical frequencies are dense in the set of all the $\le l$-step nilsequences (cf. [24, Exercise 1.6.20]; see also [15, Definition 3.4] for vertical Fourier series expansion), it suffices to prove Theorem 1.5 for the nilsequence with vertical frequency.

To prove Theorem 1.5, we will use the following estimate that first appeared in the work of Q. Chu [13] for the case $k=2$. We will show that there is a similar estimate for any number of transformations. The arguments presented here are analogous to that of the cited reference. This lemma will be useful as we apply van der Corput's lemma to the averages in (3) for k times, we will take multiplicative derivative of the $k+1$-step nilsequence for k times, which gives us a one-step nilsequence.

Lemma 4.1 ([13, Lemma 3.1]). *Suppose $(Y,\nu,S_1,\dots,S_k)$ is a system with commuting measure-preserving transformations $S_1,\dots,S_k$, and $g_0,g_1,\dots,g_k\in L^{\infty}(\nu)$. Let*

$$I_k(n)=\int g_0\prod_{i=1}^{k}g_k\circ S_k^n\,d\nu.$$

Then for any $t \in \mathbb{R}$, we have

$$\limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=0}^{N-1} e(nt) I_k(n) \right| \le |||g_0|||_{S_1,S_2,\dots,S_k}. \tag{10}$$

Proof. We can rewrite the integral $I_k(n)$ so that

$$I_k(n) = \int g_0 \circ S_1^{-n} \cdot g_1 \cdot \prod_{i=2}^{k} g_i \circ (S_i S_1^{-1})^n \, d\nu. \tag{11}$$

If $t = 0$, then (10) follows directly from [19, Propositon 1]. If $t \neq 0$, we apply the triangle inequality and the Cauchy-Schwarz inequality to the left-hand side of (10) to obtain

$$\limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=0}^{N-1} e(nt) I_k(n) \right| \le \left(\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} e(nt) g_0 \circ S_1^{-n} \prod_{i=2}^{k} g_i \circ (S_i S_1^{-1})^n \right\|_{L^2(\nu)}^2 \right)^{1/2}. \tag{12}$$

We apply van der Corput's lemma to the lim sup of the right hand side to obtain

$$\begin{aligned} &\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} e(nt) g_0 \circ S_1^{-n} \prod_{i=2}^{k} g_i \circ (S_i S_1^{-1})^n \right\|_{L^2(\nu)}^2 \\ &\le \limsup_{H\to\infty} \frac{1}{H} \sum_{h=0}^{H-1} \limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=0}^{N-1} \int (g_0 \cdot g_0 \circ S_1^{-h}) \circ S_1^{-n} \prod_{i=2}^{k} (g_i \cdot g_i \circ (S_i S_1^{-1})^h) \circ (S_i S_1^{-1})^n \, d\nu \right|. \end{aligned}$$

Since S_1 and S_2 are measure-preserving transformations, the right-hand side of the last inequality can be bounded above by

$$\limsup_{H\to\infty} \frac{1}{H} \sum_{h=0}^{H-1} \limsup_{N\to\infty} \int \left| (g_2 \cdot g_2 \circ (S_2 S_1^{-1})^h) \frac{1}{N} \sum_{n=0}^{N-1} \int (g_0 \cdot g_0 \circ S_1^{-h}) \circ S_2^{-n} \prod_{i=3}^{k} (g_i \cdot g_i \circ (S_i S_1^{-1})^h) \circ (S_i S_2^{-1})^n \right| d\nu\,,$$

so by Hölder's inequality, we have

$$\begin{aligned} &\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} e(nt) g_0 \circ S_1^{-n} \prod_{i=2}^{k} g_i \circ (S_i S_1^{-1})^n \right\|_{L^2(\nu)}^2 \\ &\le \limsup_{H\to\infty} \frac{1}{H} \sum_{h=0}^{H-1} \limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} (g_0 \cdot g_0 \circ S_1^{-h}) \circ S_2^{-n} \prod_{i=3}^{k} (g_i \cdot g_i \circ (S_i S_1^{-1})^h) \circ (S_i S_2^{-1})^n \right\|_{L^2(\nu)}, \end{aligned}$$

and by the estimate (6), the Cauchy-Schwarz inequality, and the limit formula for the box seminorm (4), we have

$$\begin{aligned} &\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} e(nt) g_0 \circ S_1^{-n} \prod_{i=2}^{k} g_i \circ (S_i S_1^{-1})^n \right\|_{L^2(\nu)}^2 \\ &\le \limsup_{H\to\infty} \frac{1}{H} \sum_{h=0}^{H-1} \left|\left|\left| g_0 \cdot g_0 \circ S_1^{-h} \right|\right|\right|_{S_2,\dots S_k} \le \left(\lim_{H\to\infty} \frac{1}{H} \sum_{h=0}^{H-1} \left|\left|\left| g_0 \cdot g_0 \circ S_1^{-h} \right|\right|\right|_{S_2,\dots S_k}^{2^{k-1}} \right)^{2^{-(k-1)}} \\ &= |||g_0|||_{S_1^{-1},S_2,\dots,S_k}^2. \end{aligned}$$

Note that, by the construction of the box seminorm, we have $|||g_0|||_{S_1^{-1},S_2,\ldots,S_k} = |||g_0|||_{S_1,S_2,\ldots,S_k}$. By the inequality (12), the claim holds. □

From this lemma, we can immediately deduce that

$$\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=0}^{N-1} e(nt)I_k(n)\right| \leq |||g_1|||_{S_1,S_2S_1^{-1},\ldots S_kS_1^{-1}} \tag{13}$$

where $I_k(n)$ is in the form of (11).

4.1. **Proof for the case** $k=2$**.** For a pedagogical purpose, we will prove Theorem 1.5 for the case $k=2$. The general case (i.e. for any $k\in\mathbb{N}$) is proved in §4.2, but the arguments are similar to that of the ones presented in here (although the notations presented here are simpler).

Proof of Theorem 1.5 *for the case* $k=2$. In this case, we assume that $f_1, f_2 \in \mathcal{Z}_3(T)$, so we know that the sequence $(f_1(T^{an}x)f_2(T^{bn}x))_n$ can be approximated by a 3-step nilsequence $(a_n)_n$. We prove this for the case that $(a_n)_n$ has a vertical frequency, and use density to show that the case holds in general (cf. [24, Exercise 1.6.20]).

We first apply van der Corput's lemma to the $L^2(\nu)$-norm of the averages to obtain an upper bound

$$\begin{aligned}
&\limsup_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1} a_n g_1\circ S_1^n g_2\circ S_2^n\right\|^2_{L^2(\nu)}\\
&\leq \liminf_{H_1\to\infty}\frac{1}{H_1}\sum_{h_1=0}^{H_1-1}\left|\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\Delta_{h_1}a_n\int g_1\cdot g_1\circ S_1^{h_1}(S_1^n y)g_2\cdot g_2\circ S_2^{h_1}(S_2^n y)d\nu(y)\right|,
\end{aligned}$$

where $\Delta_{h_1}a_n := a_{n+h_1}\overline{a}_n$ denotes the multiplicative derivative of a_n with respect to h_1. Note that $\Delta_{h_1}a_n$ is a 2-step nilsequence by [24, Lemma 1.6.13]. By applying the Cauchy-Schwarz inequality, the lim inf above is bounded above by

$$\liminf_{H_1\to\infty}\frac{1}{H_1}\sum_{h_1=0}^{H_1-1}\limsup_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}\Delta_{h_1}a_n g_1\cdot g_1\circ S_1^{h_1}(S_1^n y)g_2\cdot g_2\circ S_2^{h_1}(S_2^n y)\right\|^2_{L^2(\nu)},$$

so we again apply van der Corput's lemma to the L^2-norm above to obtain the upper estimate of

$$\liminf_{H_1\to\infty}\left(\frac{1}{H_1}\sum_{h_1=0}^{H_1-1}\liminf_{H_2\to\infty}\frac{1}{H_2}\sum_{h_2=0}^{H_2-1}\left|\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\Delta_{h_1}\Delta_{h_2}a_n\int G_{1,h_1}\cdot G_{1,h_1}\circ S_1^{h_2}(S_1^n y)G_{2,h_1}\cdot G_{2,h_1}\circ S_2^{h_2}(S_2^n y)d\nu(y)\right|\right)^{1/2},$$

where $G_{i,h_1} = g_i\cdot g_i\circ S_i^{h_1}$ for $i=1,2$. Because $\Delta_{h_1}\Delta_{h_2}a_n$ is a one-step nilsequence for each positive integers h_1 and h_2, which implies that it is a constant multiple of the exponential $e(tn)$ for some $t\in\mathbb{T}$, we can investigate this $\limsup_{N\to\infty}$ by looking at the behavior of

$$\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=0}^{N-1} e(nt)\int G_{1,h_1}\cdot G_{1,h_1}\circ S_1^{h_2}(S_1^n y)G_{2,h_1}\cdot G_{2,h_1}\circ S_2^{h_2}(S_2^n y)d\nu(y)\right|.$$

By (13), the above $\limsup$ is bounded above by $\left|\left|\left| G_{1,h_1} \cdot G_{1,h_1} \circ S_1^{h_2} \right|\right|\right|_{S_1, S_1 S_2^{-1}}$, where $|||\cdot|||$ here is the box seminorm. Hence, using the limit formula (4), the original average is bounded above by

$$\liminf_{H_1\to\infty} \left(\frac{1}{H_1} \sum_{h_1=0}^{H_1-1} \liminf_{H_2\to\infty} \frac{1}{H_2} \sum_{h_2=0}^{H_2-1} \min_{i=1,2} \left|\left|\left| G_{i,h_1} \cdot G_{i,h_2} \circ S_i^{h_2} \right|\right|\right| \right)^{1/2}$$
$$\leq \liminf_{H_1\to\infty} \left(\frac{1}{H_1} \sum_{h_1=0}^{H_1-1} \left|\left|\left| g_1 \cdot g_1 \circ S_1^{h_1} \right|\right|\right|^2 \right)^{1/2} = |||g_1|||^2_{S_1,S_1,S_1,S_1S_2^{-1}}$$

This shows that Theorem 1.5 holds for $k = 2$. □

4.2. Proof for general k.

Proof of Theorem 1.5 *for any* $k \geq 2$. As in the proof for the case $k = 2$, we assume that $f_1, f_2 \in \mathcal{Z}_{k+1}(T)$, and the sequence $(f_1(T^{an}x)f_2(T^{bn}x))_n$ is approximated by a $k+1$-step nilsequence with vertical frequency (a_n). We let $\vec{h}(j) = (h_1, h_2, \ldots, h_j) \in \mathbb{N}^j$, and for each i and j, we recursively define (on j) so that

$$G_{i,\vec{h}(1)} = g_i \cdot g_i \circ S_i^{h_1},\ G_{i,\vec{h}(2)} = G_{i,\vec{h}(1)} \cdot G_{i,\vec{h}(1)} \circ S_i^{h_2},\ \ldots,\ G_{i,\vec{h}(j)} = G_{i,\vec{h}(j-1)} \cdot G_{i,\vec{h}(j-1)} \circ S_i^{h_j}.$$

With these notations in mind, we apply van der Corput's lemma to obtain

$$\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} a_n \prod_{i=1}^k g_i \circ S_i^n \right\|^2_{L^2(\nu)}$$
$$\leq \limsup_{H_1\to\infty} \frac{1}{H_1} \sum_{h_1=0}^{H_1-1} \limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n=0}^{N-1} \Delta_{h_1} a_n \int \prod_{i=1}^k G_{i,\vec{h}(1)} \circ S_i^n \, d\nu \right|.$$

By applying the Cauchy-Schwarz (after pushing the averages and the absolute value inside the integral), we obtain

$$\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} a_n \prod_{i=1}^k g_i \circ S_i^n \right\|^2_{L^2(\nu)}$$
$$\leq \limsup_{H_1\to\infty} \left(\frac{1}{H_1} \sum_{h_1=0}^{H_1-1} \limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} \Delta_{h_1} a_n \prod_{i=1}^k G_{i,\vec{h}(1)} \circ S_i^n \right\|^2_{L^2(\nu)} \right)^{1/2}$$

And notice that we can apply this process of van der Corput's lemma and the Cauchy-Schwarz inequality again to the L^2-norm on the right hand of this inequality. We repeat this process for $k-1$ more times to obtain

$$\limsup_{N\to\infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} a_n \prod_{i=1}^k g_i \circ S_i^n \right\|^2_{L^2(\nu)}$$
$$\leq \limsup_{H_1\to\infty} \left(\frac{1}{H_1} \sum_{h_1=0}^{H_1-1} \limsup_{H_2\to\infty} \frac{1}{H_2} \sum_{h_2=0}^{H_2-1} \cdots \limsup_{H_k\to\infty} \frac{1}{H_k} \sum_{h_k=0}^{H_k-1} \right.$$

$$\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=0}^{N-1}\Delta_{h_k}\Delta_{h_{k-1}}\cdots\Delta_{h_2}\Delta_{h_1}a_n\int\prod_{i=1}^{k}G_{i,\vec{h}(k)}\circ S_i^n\,d\nu\right|\Bigg)^{2^{-(k+1)}}.$$

Since $\Delta_{h_k}\Delta_{h_{k-1}}\cdots\Delta_{h_2}\Delta_{h_1}a_n$ is a one-step nilsequence, we can apply Lemma 4.1 to show that

$$\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=0}^{N-1}\Delta_{h_k}\Delta_{h_{k-1}}\cdots\Delta_{h_2}\Delta_{h_1}a_n\int\prod_{i=1}^{k}G_{i,\vec{h}(k)}\circ S_i^n\,d\nu\right|\leq\left|\!\left|\!\left|G_{1,\vec{h}(k)}\right|\!\right|\!\right|_{\circ}$$

where $|\!|\!|\cdot|\!|\!|_{\circ}$ is the seminorm associated to the transformations $S_1, S_2S_1^{-1},\ldots,S_kS_1^{-1}$. Hence, we would have

$$\begin{aligned}&\limsup_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}a_n\prod_{i=1}^{k}g_i\circ S_i^n\right\|_{L^2(\nu)}^2\\ &\leq\limsup_{H_1\to\infty}\left(\frac{1}{H_1}\sum_{h_1=0}^{H_1-1}\limsup_{H_1\to\infty}\frac{1}{H_2}\sum_{h_2=0}^{H_2-1}\cdots\limsup_{H_k\to\infty}\frac{1}{H_k}\sum_{h_k=0}^{H_k-1}\left|\!\left|\!\left|G_{1,\vec{h}(k)}\right|\!\right|\!\right|_{\circ}\right)^{2^{-(k+1)}}.\end{aligned}$$

When we apply the Cauchy-Schwarz inequality and the limit formula (4), the upper bound in the above inequality becomes

$$\begin{aligned}&\limsup_{H_1\to\infty}\left(\frac{1}{H_1}\sum_{h_1=0}^{H_1-1}\limsup_{H_1\to\infty}\frac{1}{H_2}\sum_{h_2=0}^{H_2-1}\cdots\limsup_{H_k\to\infty}\left(\frac{1}{H_k}\sum_{h_k=0}^{H_k-1}\left|\!\left|\!\left|G_{1,\vec{h}(k)}\right|\!\right|\!\right|_{\circ}^{2^k}\right)^{2^{-k}}\right)^{2^{-(k+1)}}\\ &=\limsup_{H_1\to\infty}\left(\frac{1}{H_1}\sum_{h_1=0}^{H_1-1}\limsup_{H_1\to\infty}\frac{1}{H_2}\sum_{h_2=0}^{H_2-1}\cdots\limsup_{H_{k-1}\to\infty}\frac{1}{H_{k-1}}\sum_{h_{k-1}=0}^{H_{k-1}-1}\left|\!\left|\!\left|G_{1,\vec{h}(k-1)}\right|\!\right|\!\right|_{S_1,\circ}^{2}\right)^{2^{-(k+1)}}\end{aligned}$$

By iterating this procedure, we will obtain

$$\limsup_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}a_n\prod_{i=1}^{k}g_i\circ S_i^n\right\|_{L^2(\nu)}^2\leq[\![\![g_1]\!]\!]_{1,k}^2,$$

and this completes the proof. □

5. Proof of Theorem 1.3

We are now ready to prove the main theorem.

Proof of Theorem 1.3. To prove the main result, we will first obtain a set of full measure $X_k\subset X$ for each $k\in\mathbb{N}$ such that for any $x\in X_k$, $a,b\in\mathbb{Z}$, and for any other measure-preserving system with k transformations $(Y,\nu,S_1,\ldots,S_k)$ with any $g_1,\ldots,g_k\in L^\infty(\nu)$, the averages

$$\frac{1}{N}\sum_{n=0}^{N-1}f_1(T^{an}x)f_2(T^{bn}x)\prod_{i=1}^{k}g_i\circ S_i^n$$

converge in $L^2(\nu)$. We will proceed proving this claim by induction on k.

The base case $k=1$ follows immediately from the double recurrence Wiener-Wintner theorem [5]. Now assume that the theorem holds for $k-1$ so that there exists a set of full measure X_{k-1} for which the theorem holds for $k-1$ measure-preserving transformations $S_1,\dots,S_{k-1}$ and functions $g_1,\dots,g_{k-1}$. To show that the theorem holds for k, we first consider the system

$$(Y,\nu,S_1,S_2,\dots,S_k,\underbrace{\mathrm{Id},\dots,\mathrm{Id}}_{k\text{ terms}}),$$

where Id denotes the identity transformation on Y. We let $U_1=S_1$, $U_i=S_iS_1^{-1}$ for $2\le i\le k$, and $U_j=S_1^{-1}$ for $k+1\le j\le 2k$, and consider the space

$$(Y^*,\nu^*,U_1^*,U_2^*,\dots,U_k^*,\underbrace{U_j^*,\,\dots,\,U_{2k}^*}_{k\text{ terms}}),$$

where the notations are described as in §2.2 i.e. $Y^*=Y^{2k}$, ν^* is the box measure associated to the transformations above, and U_i^* is the side transformation of U_i in Y^* for each $i=1,2,\dots,2k$. Note that for $2\le i\le k$, $S_i^*=U_i^*U_1^*$, and we observe that the system $(Y,\nu,S_1,S_2,\dots,S_k,\mathrm{Id},\dots,\mathrm{Id})$ is a factor of $(Y^*,\nu^*,S_1^*,\dots,S_k^*,\mathrm{Id}^*,\dots,\mathrm{Id}^*)$. Since there exists a factor map $\pi:Y^*\to Y$ such that $S_i\circ\pi=\pi\circ S_i^*$ for each i, it suffices to show that there exists a set of full measure $X_k\subset X$ such that for any $x\in X_k$ and any other measure-preserving system with commuting transformations $(Y,\nu,S_1,S_2,\dots,S_k)$, the averages

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^{an}x)f_2(T^{bn}x)\prod_{i=1}^{k} g_i^*\circ S_i^* \tag{14}$$

converge in $L^2(\nu^*)$.

We first consider the case g_1^* is $\mathcal{W}^*$-measurable, where

$$\mathcal{W}^*=\bigvee_{i=1}^{2k}\mathcal{I}(U_i^*)=\bigvee_{i=1}^{k}\mathcal{I}(U_i^*),$$

since for $k+1\le j\le 2k$, $\mathcal{I}(U_j^*)=\mathcal{I}(S_1^{*-1})=\mathcal{I}(S_1^*)=\mathcal{I}(U_1^*)$. We further consider the case

$$g_1^*=\prod_{i=1}^{k} h_i^*,\text{ where for each } h_i^*\in L^\infty(\nu^*),\, 1\le i\le k,\, h_i^*\in\mathcal{I}(U_i^*) \tag{15}$$

Then the averages in (14) can be expressed as

$$h_1^*\cdot\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^{an}x)f_2(T^{bn}x)\prod_{i=2}^{k}(g_i^*\cdot h_i^*)\circ S_i^*,$$

and by the inductive hypothesis, the averages in above converge for all $x\in X_{k-1}$ in $L^2(\nu^*)$.

Because the linear span of functions of the form of (15) is dense in $L^\infty(\nu^*,\mathcal{W}^*)$ (in $L^1(\nu^*)$-norm), the density argument tells us the averages in (14) converge for all $x\in X_{k-1}$.

To prove the inductive step, it remains to show that the claim holds for the case $\mathbb{E}(g_1^*|\mathcal{W}^*)=0$. This case can be treated by breaking into two sub-cases: The sub-case where either $\mathbb{E}(f_i|\mathcal{Z}_{k+1}(T))=0$ for

$i = 1, 2$, or the sub-case where both $f_1, f_2 \in \mathcal{Z}_{k+1}(T)$. The first sub-case is treated by Theorem 1.4, so there exists a set of full measure $X^1_{f_1,f_2}$ for which the averages converge to 0 in $L^2(\nu)$. For the second sub-case, the fact that the system Y^* is magic [19, Theorem 2] implies that $\left|\left|\left|g_1^*\right|\right|\right|^* = 0$, where $|||\cdot|||^*$ is the box seminorm associated to the transformations $U_1^*, U_2^*, \ldots, U_{2k}^*$, or in other names,

$$S_1^*, S_2^* S_1^{*-1}, \ldots, S_k^* S_1^{*-1}, \underbrace{S_1^{*-1}, \ldots, S_1^{*-1}}_{k \text{ times}}.$$

By the construction of the box seminorm, we know that

$$|||g_1^*|||^* = [[[g_1^*]]]^*_{1,k}\,,$$

where $[[[\cdot]]]^*_{1,k}$ is the seminorm seen in Theorem 1.5, associated to the transformations

$$S_1^*, S_2^* S_1^{*-1}, \ldots, S_k^* S_1^{*-1}, \underbrace{S_1^*, \ldots, S_1^*}_{k \text{ times}}$$

(this follows from the fact that the seminorm remains unchanged if S_1^{*-1} is replaced by S_1^*). By the fact that the sequence $a_n = f_1(T^{an}x)f_2(T^{bn}x)$ can be approximated by a $k+1$-step nilsequence, we apply Theorem 1.5 to find a set of full measure $X^2_{f_1,f_2}$ for which the averages converge to 0 in $L^2(\nu)$. Take $X_k = X_{k-1} \cap X^1_{f_1,f_2} \cap X^2_{f_1,f_2}$, and we complete the inductive step.

To conclude the proof, we set $X_{f_1,f_2} = \bigcap_{k=1}^\infty X_k$, and we obtain the desired set of full measure for which the theorem holds. □

References

[1] I. Assani. Multiple recurrence and almost sure convergence for weakly mixing dynamical systems. *Israel J. Math.*, 103:111–124, 1998.

[2] I. Assani. Multiple return times theorems for weakly mixing systems. *Ann. Inst. Henri Poincaré, Probabilités et Statistiques*, 36:153–165, 2000.

[3] I. Assani. Pointwise convergence of ergodic averages along cubes. *J. Analyse Math.*, 110:241–269, 2010.

[4] I. Assani. Pointwise double recurrence and nilsequences. Preprint. Available from arXiv:1504.05732, 2015.

[5] I. Assani, D. Duncan, and R. Moore. Pointwise characteristic factors for Wiener-Wintner double recurrence theorem. *Ergod. Th. and Dynam. Sys.*, 2015. Available on CJO 2015 doi:10.1017/etds.2014.99.

[6] I. Assani and R. Moore. Extension of Wiener-Wintner double recurrence theorem to polynomials. To appear in *Journal d'Analyse Mathematique*, arXiv:1409.0463, 2015.

[7] I. Assani and R. Moore. A good universal weight for nonconventional ergodic averages in norm. To appear in *Ergod. Th. and Dynam. Sys.*, available on arXiv:1503.08863., 2015.

[8] I. Assani and K. Presser. Survey of the return times theorem. In *Ergodic Theory and Dynamical Systems: Proceedings of the Ergodic Theory Workshops at University of North Carolina at Chapel Hill, 2011-2012*, De Gruyter Proc. Math., pages 19–58. De Gruyter, Berlin, 2014.

[9] T. Austin. On the norm convergence of non-conventional ergodic averages. *Ergodic Theory Dynam. Systems*, 30(2):321–338, 2010.

[10] J. Bourgain. Return time sequences of dynamical systems. Unpublished preprint, 1988.

[11] J. Bourgain. Double recurrence and almost sure convergence. *J. reine angew. Math.*, 404:140–161, 1990.

[12] J. Bourgain, H. Furstenberg, Y. Katznelson, and D. Ornstein. Appendix on return-time sequences. *Publ. Math. Inst. Hautes Études Sci*, 69:42–45, 1989.

[13] Q. Chu. Multiple recurrence for two commuting transformations. *Ergodic Theory Dynam. Systems*, 31(3):771–792, 2011.

[14] D. Duncan. *A Wiener-Wintner Double Recurrence Theorem*. PhD thesis, The University of North Carolina at Chapel Hill, 2001. Advisor: I. Assani.

[15] T. Eisner and P. Zorin-Kranich. Uniformity in the Wiener-Wintner theorem for nilsequences. *Discrete and Continuous Dynamical Systems*, 33(8):3497–3516, 2013.

[16] N. Frantzikinakis. Multiple correlation sequences and nilsequences. arXiv:1407.0631, to appear in *Inventiones Mathematicae*.

[17] W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11:465–588, 2001.

[18] B. Green and T. Tao. An inverse theorem for the Gowers U^3 norm. *Proc. Edinburgh Math. Soc*, 51(1):73–153, 2008.

[19] B. Host. Ergodic seminorms for commuting transformations and applications. *Studia Math.*, 195(1):31–49, 2009.

[20] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math.*, 161:387–488, 2005.

[21] B. Host and B. Kra. Uniformity seminorms on ℓ^∞ and applications. *J. Anal. Math*, 108:219–276, 2009.

[22] W. Huang, S. Shao, and X. Ye. Pointwise convergence of multiple ergodic averages and strictly ergodic models. Preprint, arXiv1406.5930, 2014.

[23] T. Tao. Norm convergence for multiple ergodic averages for commuting transformations. *Ergod. Th. and Dynam Sys.*, 28:657–688, 2008.

[24] T. Tao. *Higher Order Fourier Analysis*. American Mathematical Society, 2012.

[25] H. Towsner. Convergence of diagonal ergodic averages. *Ergod. Th. and Dynam. Sys.*, 29(4):1309–1326, 2009.

[26] T. Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.*, 20(1):53–97, 2006.

[27] P. Zorin-Kranich. A nilsequence Wiener-Wintner theorem for bilinear ergodic averages. Preprint. Available from arXiv:1504.04647, 2015.

Department of Mathematics, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

E-mail address: `assani@math.unc.edu`

URL: `http://www.unc.edu/math/Faculty/assani/`

Department of Mathematics, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

E-mail address: `ryom@live.unc.edu`

URL: `http://ryom.web.unc.edu`